AD-781 035

# OPTIMAL ASSIGNMENT OF AIR FORCE PILOTS

Robert E. Miller

Air Force Human Resources Laboratory
Brooks Air Force Base, Texas

February 1974

# NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed and cleared for open publication and/or public release by the appropriate Office of Information (OI) in accordance with AFR 190-17 and DoDD 5230.9. There is no objection to unlimited distribution of this report to the public at large, or by DDC to the National Technical Information Service (NTIS).

This final report was submitted by Personnel Research Division, Air Force Human Resources Laboratory, Lackland Air Force Base, Texas 78236, under project 7719, with HQ Air Force Human Resources Laboratory, Brooks Air Force Base, Texas 78235.

This technical report has been reviewed and is approved.

LELAND D. BROKAW, Chief
Personnel Research Division

Approved for publication.

HAROLD E. FISCHER, Colonel, USAF
Commander

PES: 3740, 3908, 3966

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

# REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | |
|---|---|
| 1. REPORT NUMBER: AFHRL-TR-73-59 | 2. GOVT ACCESSION NO. |
| 3. RECIPIENT'S CATALOG NUMBER: AD 781 035 | |
| 4. TITLE (and Subtitle): OPTIMAL ASSIGNMENT OF AIR FORCE PILOTS | 5. TYPE OF REPORT & PERIOD COVERED: Final |
| | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s): Robert E. Miller | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS: Personnel Research Division, Air Force Human Resources Laboratory, Lackland Air Force Base, Texas 78236 | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 77191208 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS: HQ Air Force Human Resources Laboratory, Brooks Air Force Base, Texas 78235 | 12. REPORT DATE: February 1974 |
| | 13. NUMBER OF PAGES: 22 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | 15. SECURITY CLASS. (of this report): Unclassified |
| | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

NATIONAL TECHNICAL INFORMATION SERVICE

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Air Force Officer Qualifying Test
Automatic Interaction Detector model
biographical data
coefficient of concordance
discriminant analysis
validation
mathematical prediction
fighter pilots
peer ranking
performance evaluation
pilot assignments
transport pilots
pilot surveys
pilot training grades
reconnaissance pilots
Tactical Air Command

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study addresses the problem, posed by Tactical Air Command, of whether a pilot on completion of undergraduate pilot training can be optimally assigned to a particular type of aircraft or mission. The problem was approached by using peer ranks to identify pilots of above average competence in each of three specialties. Using only these pilots as subjects, a multiple discriminant analysis was performed to yield a system for identifying a unique assignment for each pilot. The system uses ten test scores and training grades to classify a new pilot as optimally assignable to a transport, fighter, or reconnaissance aircraft or mission. The peer ranking was found to be predictable, and pilots actually assigned in accordance with their optimal assignments were shown to be better pilots, as measured by the peer ranking, than pilots not optimally assigned.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

I

## SUMMARY


### Problem

Pilots are trained for their specific assignments. Yet aircraft accidents are often attributed to pilot error. These observations raise a question of whether there is an optimal assignment for a given pilot to some specific type of aircraft or mission.

### Approach

Tactical Air Command (TAC) pilots were asked to rank other pilots of their organization in terms of competence. Additional data pertaining to operational test scores, undergraduate pilot training grades, and biographical information were obtained from personnel files and an experimental questionnaire. These background variables were used to predict the peer ranking, and subsets of them were used in a multiple discriminant analysis, based only on pilots of above average competence, to develop a system for assigning a pilot to a particular specialty.

### Results

The peer ranking was found to be predictable within the specialties of transport, fighter, and reconnaissance pilot, and in all three specialties combined. An efficient set of ten test scores and training grades was identified for discriminating among above average pilots in the three specialty groups.

Among above average pilots, 58 percent are actually assigned compatibly with their optimal assignments. Among average and below average pilots, 43 percent are so assigned. Within each specialty, the optimally assigned pilots have a higher average peer ranking than the others. This is taken as evidence for the validity of the assignment system. It was found possible to make optimal assignments within TAC quotas for 85 percent of the pilots in the study and to assign all the remainder with some regard for their optimal assignments.

### Conclusions

An optimal assignment system for transport, fighter, and reconnaissance pilot specialties is possible. Pilots with optimal assignments are significantly better pilots than others as determined by peer ranking.

## PREFACE

This research was requested by Tactical Air Command (TAC) under Requirement for Personnel Research 69-4, Pilot Selective Assignment Study. Research was performed by the Personnel Research Division, Air Force Human Resources Laboratory, under Project 7719, Air Force Personnel System Development on Selection, Assignment, Evaluation, Quality Control, Retention, Promotion, and Utilization; Task 771912, Selection and Classification Instruments for Officer Personnel Programs.

Appreciation is expressed to Lt General Momyer, Commander of TAC, who took a personal interest in the initiation of this study; to Col Carkeet and Lt Col Van Huss of Hq TAC, who made available the basic data and offered useful suggestions on the survey of TAC pilots; and to Maj Cline of Hq USAF (AFPRMS), who made available the data from Project Sabrewings and assisted in the analysis.

Preceding page blank

## TABLE OF CONTENTS



## LIST OF ILLUSTRATIONS



## LIST OF TABLES





**Preceding page blank**

List of Tables (*Continued*)

# OPTIMAL ASSIGNMENT OF AIR FORCE PILOTS

## I. INTRODUCTION

Tactical Air Command (TAC) has observed that the primary cause of many aircraft accidents is pilot error. Frequently the error arises from inability to maintain control of the aircraft or situation during normal operations. This occurs even though pilots receive training for their specific assignments.

Since the demands on each pilot are in part peculiar to his type of aircraft and mission, there may be an optimal assignment where he will perform better than in any other. Current assignment policies do not have a research basis which supports this optimization. An optimized system, however, should result not only in improved flying safety, but also in greater mission effectiveness, greater job satisfaction, and more efficient training programs. TAC has pointed out that these benefits should accrue to any command having operational aircraft.

## II. HISTORICAL BACKGROUND

Studies which attempt to predict pilot performance are limited almost wholly to student pilots. The first large scale effort was initiated in World War II as a way of meeting the problem of excessive attrition from pilot training. One resulting product was the Aircrew Classification Battery, the first form of which became operational in February 1942. This was a battery of aptitude tests which yielded stanine scores predictive of performance as a pilot, navigator, or bombardier.

When the battery was revised in September 1944, the pilot score was replaced by separate scores for fighter and bomber pilots. The rationale was that, since training programs for the two specialties were different, the aptitude requirements should also be different. This assumption was supported by job analyses. The two scores were actually used to classify candidates, but, because there was a reduced need for pilots, few of those classified were ever trained. Hence the fighter and bomber scores were not fully evaluated. It is known that a correlation of .90 existed between them (DuBois, 1947).

After World War II, pilot training became a one-track effort in which all students were trained on the same type of aircraft. It was expected that students who completed the training and were assigned to other types of aircraft could be cross-trained. In this setting there was little interest in the differential classification of candidates. The single pilot score was reinstated as the sole aptitude qualification.

The Air Force concurred in the widely accepted philosophy that aptitude testing is for the prediction of performance in training, and that performance on the job is a function of the training. There are few studies which relate aptitudes directly to performance as manifested in accidents. One such study was accomplished as part of the World War II research in which over one thousand students were admitted to pilot training without regard for aptitudes. Twenty accidents involving training aircraft occurred in this group. Fifteen of these accidents, including all the fatalities, involved students with pilot aptitude stanines of 6 or below. A correlation of .12 is reported between the pilot stanine and accidents among 3,500 pilots in operational training (Flanagan, 1948).

## III. APPROACH

The problem posed by TAC does not require an analysis in terms of accidents. Accidents can be used actuarially in relation to test scores, but they are not highly sensitive to individual differences in performance because most pilots do not have accidents. Moreover, there is an uncertain sensitivity in such measures as the Officer Effectiveness Report (OER), Combat Crew Training School (CCTS) grades, Standarization-Evaluation Board proceedings, and decorations for flying. The performance measure finally adopted was a peer ranking in terms of pilot competence.

The purpose of the performance measure was to identify above average pilots in each specialty studied. These pilots then became the subjects for the development of the actual assignment system. The system was based on operational test scores, undergraduate pilot training (UPT) grades, and biographical data. These variables were subjected to a multiple discriminant analysis to find one or more linear functions which maximize the differences among the specialty groups. Under this system, when a new pilot is to be assigned, his values on the variables are used to compute a discriminant score which will uniquely categorize him in one of the specialty groups. Since above average competence is made a part of the definition of the groups, his categorization will be in that group where he most closely resembles an above average pilot.

Predictability of the peer ranking is of interest largely because there are no data on whether performance of operational pilots can be predicted within homogeneous specialties. The problem of predicting the peer ranking was therefore included in the study, both in terms of zero order correlations and the Automatic Interaction Detector (AID) model (Koplyay, Gott, & Elton, 1973). The AID model yields multiple correlations with the peer ranking based on the variables and all possible interactions among them. This model was used in lieu of the original University of Michigan model (Sonquist & Morgan, 1964).

## IV. THE PREDICTORS

A restriction imposed by TAC is that the optimal assignment of a pilot be made from variables whose values are known by the end of UPT. Within this limitation, there are several possible sets of discriminating or predictive variables. The sets used in this study were:

1. Percentile scores on the Air Force Officer Qualifying Test (AFOQT). One of these scores was designed specifically to predict performance of student pilots and is used operationally in selecting candidates for UPT. Two other scores on this test have been shown to predict some facet of student pilot performance (Miller, 1969).

2. Grades and other records of experience in UPT. It is assumed that if performance on the job is a function of training, measures of performance in training should be predictive of success on the job. Research on student pilots in the Navy (Peterson, Booth, Lane, & Ambler, 1967) and civilian life (Hulin & Alvares, 1971) support this assumption. Because several UPT grading systems have been used, it was necessary to convert all grades to the current system. This was done through tables from the Historical Section of Air Training Command.

3. Biographical data obtained both from personnel files and an experimental survey of pilots assigned to TAC. Biographical data have been used to delineate background factors which characterize fighter pilots and to distinguish fighter pilots from bomber pilots (Torrance, 1954).

Two other variables were used for the special purpose of studying relationships among possible performance measures. These variables are the mean weighted OER and the CCTS grade. Because they can not be considered available by the end of UPT, they have no role in determining optimal pilot assignments.

Some of the variables are continuous in that they can take any numerical value within a given range. Others are dichotomous. For these, a value of one is assigned to individuals possessing the characteristic described by the variable, and a value of zero is assigned to all others.

## V. THE PERFORMANCE MEASURE

The peer rating in its various forms has frequently been used in Air Force training programs, either as a predictor or as a performance criterion. The Cadet Effectiveness Rating at the Air Force Academy is an example. Peer ratings are widely used with reasonably good results where no other acceptable measure exists.

Because there was doubt about what specific behavior is crucial in pilot performance, it seemed best to use a global type rating in which each pilot is asked to rank every other pilot in his organization in terms of over-all competence as a pilot. The intent was to have each pilot assign a rank of one to the best pilot in his organization and to rank the other pilots in order. It was not desired that he rank himself. It was believed that ranking within organizations would tend to prevent the ranking of pilots by others having a different specialty. Identification of the organizations, at the squadron level or equivalent where possible, was left to TAC.

The set of ranks which were the basic performance measure do not constitute a true scale of measurement. To overcome this deficiency, the ranks assigned to each pilot were averaged and the results converted to positions on a scale of normalized standard scores having a theoretical mean of 50 and standard deviation of 10. This conversion involves the questionable assumption that pilot competence is the same in all organizations. The consequence for the study, should the assumption be incorrect, is that the performance measure will be difficult to predict from any set of variables. As a final step in scaling the rankings, the ends of the scale were reversed so high values would be associated with low ranks, representing the pilots judged most competent.

## VI. THE PILOT SURVEY

TAC provided a listing by organization of all pilots in the command occupying cockpit assignments as of 23 September 1969, together with listings of projected gains and losses. These listings were used to compile a complete enumeration of the TAC pilot population as of 1 November 1969. Not all pilots in the enumeration were used in the study. It was suggested by the TAC study monitor that the task of evaluating competence would be distasteful to pilots. On his recommendation, pilots above the rank of major and all unit commanders were excluded. Majors in large organizations were also excluded to keep within bounds the number of pilots to be ranked. Projected gains were excluded so pilots would not be asked to rank others with whom they were not well acquainted. Finally, pilots with classified assignments were excluded.

To each remaining pilot was sent, through the Test Control Officer at his base, a packet containing a set of mark-sense cards prepunched and printed with the identity of each pilot in his unit other than himself, a letter explaining the study and citing the authority for it, a biographical questionnaire known as the Background Information Survey, and a sheet of instructions. Rankings were to be made on the mark-sense cards. The completed rankings and questionnaires were to be returned directly to the research facility. An explanation of the study was sent to each Test Control Officer, and a follow-up letter was sent to each nonresponding pilot after approximately two weeks.

Returns were received for the next several months. The return rate eventually reached 63 percent, including cases who refused the ranking task but completed the questionnaire. These cases were initially included on the assumption that, if they were ranked by any of their peers, a usable set of data for them existed. Many sets of ranks had to be rejected because of idiosyncratic responses. It is likely that personal administration of the materials by an experimenter could have mitigated this problem.

## VII. THE SAMPLE

Data from all usable returns were key punched, placed on magnetic tape, and matched by computer against officer personnel files to pick up AFOQT scores and additional biographical data. Most matching operations resulted in a loss of some cases for failure to effect a match. AFOQT scores not found by matching were retrieved manually from files at the Air Force Military Personnel Center. UPT grades were extracted manually from Summary Records of Flying Training. Some variables were missing from matching files in enough cases to warrant discarding the variables. Since only 60 cases had data complete in every respect, it was decided to use mean or estimated values in place of missing data when there were relatively few cases affected. The resulting distortion has the effect of restricting variability in the sample and discouraging the discovery of relationships which do not actually exist.

Many cases were lost in the operation of classifying pilots by specialty. This was done according to the Duty Air Force Specialty Code (DAFSC). In a force as diverse as TAC there are many pilot DAFSCs, some of them occurring only in small numbers. DAFSCs having few cases in the sample were combined with similar DAFSCs where the Officer Classification Manual implied that this was appropriate. Otherwise they were dropped from the study.

Cases still in the study at this point constituted the final sample. There were 784 such cases, divided into three groups on the basis of DAFSC. One group consisted of 230 transport pilots with DAFSCs 1045, 1051, and 1055. Another group consisted of 485 fighter pilots with DAFSCs 1111 and 1115. A third group consisted of 69 reconnaissance (recce) pilots with DAFSCs 1321 and 1325. The recce group is of marginal size for analysis, but it was nevertheless retained in the study.

## VIII. EVALUATION OF THE SAMPLE

The exclusions of certain pilots from the study were nonrandom. An attempt was therefore made to determine the degree to which the final sample is representative of the TAC population. There were few bases for comparison, but it was possible to compare the sample and population in terms of the distributions of DAFSCs and military ranks. These comparisons are presented in Table 1.

The first section of the table compares the proportions of the sample falling in each of the three DAFSC groups with the proportions of the population in the same DAFSC groups, ignoring other DAFSCs in the population. Thus, the total of each column of proportions is 1.00 except for rounding errors. There is good agreement between the sample and population.

*Table 1.* Comparison of TAC Sample and Population by DAFSC Group and Rank

*(Sample N = 784)*

| DAFSC Group | Proportion in Sample | Proportion in Population | Rank | Proportion in Sample | Proportion in Population |
|---|---|---|---|---|---|
| Transport | .29 | .34 | Major | .22 | .39 |
| Fighter | .62 | .59 | Captain | .57 | .43 |
| Recce | .09 | .08 | 1st Lieutenant | .21 | .14 |
| | | | 2nd Lieutenant | .00 | .04 |
| Total | 1.00 | 1.01 | Total | 1.00 | 1.00 |

The second section of the table compares the proportions of the sample having ranks from 2nd lieutenant through major with the proportions of the population having the same ranks, ignoring other ranks in the population. The agreement here is not as good. Except for the 2nd lieutenants, there is an overrepresentation of junior officers in the sample. Also, there is an underrepresentation of majors as a direct consequence of one of the exclusion operations. This is probably not a serious deficiency because the results of the study are intended for application to new UPT graduates, nearly all of whom are junior officers. There were three 2nd lieutenants in the sample, amounting to a proportion of less than .005.

A further question arises of how much of the TAC population is represented by the three DAFSC groups and the ranks through major. The listing of all TAC pilots with cockpit assignments shows that 93 percent fall in one or another of the DAFSC groups. Also, 82 percent fall in one or another of the DAFSC groups and have a rank of major or below. It is important to note that the sample represents DAFSC groups associated with fighter and heavy multi-engine types of aircraft. This is a fundamental aircraft classification.

## IX. THE SUBSAMPLES

The three DAFSC groups were each split randomly into subsamples arbitrarily designated A and B. The number of cases in each subsample A was identical, or nearly so, with the number in the corresponding subsample B. These random splits were made so the predictive portion of the analysis could be performed on subsample A and applied to subsample B, thus allowing for cross-validation.

Subsample A for the combined DAFSC groups contained 392 cases. In Table 2 the characteristics of this subsample are described in terms of the mean and standard deviation of each variable. Somewhat different means and standard deviations are found in the DAFSC groups separately, but many of the differences are too small to be interpretable. The data for subsample B are assumed to be essentially identical with those for subsample A.

No use was made of the distinction between subsamples A and B in the discriminant analysis. Rather, each DAFSC group was split at its own mean peer ranking into an above average group of pilots and a group of average and below average pilots. The analysis performed on the above average group was applied to all pilots regardless of their group membership.

One additional subsample of some importance was defined. This cuts across other subsamples and consists of 92 fighter pilots who were found to have been subjects in a Headquarters USAF study known as Project Sabrewings. These cases were discovered when the Sabrewings files were examined for possible additional variables. Mean weighted OERs and CCTS grades for the 92 cases were extracted to study their relationships with the peer rankings and with each other. The correlations ranged from .01 to .14, and none were statistically significant.

## X. EVALUATION OF THE PERFORMANCE MEASURE

TAC pilots were instructed not to collaborate in performing the ranking task. However, if a pilot is assigned some ranks indicating a high level of competence and some indicating a low level, no

*Table 2.* **Characteristics of Combined DAFSC Groups**
**($N$=392)**

| Variable | Mean | SD |
|---|---|---|
| **AFOQT Variables** | | |
| AFOQT Pilot Percentile | 69.78 | 21.22 |
| AFOQT Navigator-Technical Percentile | 67.19 | 23.15 |
| AFOQT Officer Quality Percentile | 60.70 | 24.09 |
| AFOQT Verbal Percentile | 56.82 | 25.91 |
| AFOQT Quantitative Percentile | 57.54 | 25.75 |
| **UPT Variables** | | |
| Rank Cadet on Entry into UPT | 0.05 | 0.21 |
| Rank 2nd Lieutenant on Entry into UPT | 0.86 | 0.34 |
| Rank 1st Lieutenant on Entry into UPT | 0.04 | 0.19 |
| Rank Captain on Entry into UPT | 0.05 | 0.22 |
| T-37 Primary Aircraft | 1.00 | 0.00 |
| Principles of Flight, Primary | 87.52 | 3.55 |
| Aviation Physiology, Primary | 87.38 | 3.89 |
| Engineering and Systems, Primary | 86.97 | 3.57 |
| Flight Instruments, Primary | 87.65 | 3.89 |
| Navigation, Primary | 87.83 | 3.55 |
| Contact Flying, Primary | 87.29 | 3.35 |
| Instrument Flying, Primary | 87.90 | 3.59 |
| B-25 Basic Aircraft | 0.01 | 0.07 |
| T-33 Basic Aircraft | 0.23 | 0.42 |
| T-38 Basic Aircraft | 0.77 | 0.42 |
| Aviation Physiology, Basic | 86.06 | 2.40 |
| Engineering and Systems, Basic | 87.27 | 4.53 |
| Instrument Procedures, Basic | 87.21 | 4.11 |
| Academic Average, Basic | 86.89 | 3.63 |
| Navigation, Basic | 87.44 | 4.51 |
| Contact Flying, Basic | 86.93 | 3.89 |
| Instrument Flying, Basic | 87.67 | 3.66 |
| Formation Flying, Basic | 87.47 | 3.84 |
| Flying Training Grade | 86.84 | 3.42 |
| Hours Light Aircraft in UPT | 24.00 | 26.22 |
| Hours Jet Aircraft in UPT | 195.74 | 68.67 |
| **Biographical Variables** | | |
| Source of Commission AFROTC | 0.53 | 0.50 |
| Source of Commission Officer Training School | 0.22 | 0.41 |
| Source of Commission Service Academy | 0.05 | 0.21 |
| Source of Commission Officer Candidate School | 0.02 | 0.15 |
| Source of Commission Other | 0.19 | 0.39 |
| Flying Experience Prior to UPT | 0.55 | 0.50 |
| Professional or Managerial Occupation of Parent | 0.53 | 0.50 |
| Age at End of UPT | 23.71 | 4.18 |
| Married by End of UPT | 0.53 | 0.50 |
| Number of Dependents at End of UPT | 1.05 | 1.21 |
| Years of Education Beyond High School at End of UPT | 3.73 | 1.41 |
| Contemplated Career in Military Aviation at End of UPT | 0.79 | 0.40 |
| Favorable Family Attitude Toward Career in Military Aviation at End of UPT | 0.49 | 0.50 |
| Favorable Attitude Toward First Assignment after UPT | 0.68 | 0.47 |
| Academic Field Business | 0.15 | 0.36 |
| Academic Field Education | 0.22 | 0.42 |
| Academic Field Engineering | 0.04 | 0.20 |
| Academic Field Mathematics | 0.06 | 0.24 |
| Academic Field Service Academy Unspecialized | 0.05 | 0.23 |
| Academic Field Social Science or Liberal Arts | 0.12 | 0.33 |
| Academic Field Other Science | 0.12 | 0.32 |
| Academic Field Other Non-Science | 0.09 | 0.29 |
| Academic Field Unspecified | 0.14 | 0.35 |
| **Other Variables** | | |
| Mean Weighted OER[1] | 7.97 | 0.45 |
| CCTS Grade[1] | 3.77 | 0.49 |
| **Performance Measure** | | |
| Scaled Peer Rank | 50.82 | 7.47 |

[1]Based on subsample of 92 cases.

useful measure of his competence has been obtained. Agreement among pilots in ranking other pilots is the fundamental reliability issue for this performance measure. Reasonable agreement is essential, and it is desirable to have a measure of it.

It was intended that the coefficient of concordance, modified to take the absence of self rankings into account, would be the measure of agreement in this study. The coefficient of concordance can take values from .00, indicating no agreement, to 1.00, indicating perfect agreement. It is applicable, however, only when each pilot ranks every other pilot in his organization. This proved to be frequently not the case.

A solution was found by computing the coefficients on the largest subset of data which was free of missing ranks from each organization. Since some data are not used, it must be assumed that the results are generalizable to all data in the study. The distribution of coefficients of concordance is shown in Table 3. The range is from .11 to 1.00, with a median value of .62. Each coefficient was evaluated by the Chi square test or by Kendall's tables, depending on the size of the organization, to determine whether the agreement it represents is greater than can be attributed to chance (Kendall, 1945). Significant results at the .05 level were achieved in 41 of 42 organizations.

*Table 3.* **Distribution of Coefficients of Concordance**

| Coefficient of Concordance | Number of Organizations |
|---|---|
| .11 - .20 | 1 |
| .21 - .30 | 1 |
| .31 - .40 | 3 |
| .41 - .50 | 8 |
| .51 - .60 | 7 |
| .61 - .70 | 10 |
| .71 - .80 | 9 |
| .81 - .90 | 2 |
| .91 - 1.00 | 1 |
| Total | 42 |

There are other indications of reasonableness in the peer rankings. It can be hypothesized that experienced pilots should generally be more highly competent than pilots at the entry level. Since the fourth digit of the specialty code indicates the level, it is possible to compute the mean ranking for each level. The peer ranking distributions at each level, together with means and standard deviations, are shown in Table 4. The difference between means is in the expected direction. It is also statistically significant, as indicated by the $t$ value of 8.60. The mean of four members of the Aerial Demonstration Squadron who were identified in the study is 52.23. This is slightly higher than the mean of all experienced pilots.

*Table 4.* **Distributions of Peer Ranks for Entry Level and Other Pilots**

| Scaled Peer Rank | Entry Level Pilots ($N$ = 234) | Other Pilots ($N$ = 550) |
|---|---|---|
| 31.50 - 37.49 | 7 | 8 |
| 37.50 - 43.49 | 57 | 69 |
| 43.50 - 49.49 | 86 | 126 |
| 49.50 - 55.49 | 61 | 156 |
| 55.50 - 61.49 | 18 | 143 |
| 61.50 - 67.49 | 5 | 37 |
| 67.50 - 73.49 | 0 | 8 |
| 73.50 - 79.49 | 0 | 3 |
| Mean | 47.64 | 52.14 |
| Standard Deviation | 6.31 | 7.51 |
| t | 8.60 | |

## XI. PREDICTION OF THE PERFORMANCE MEASURE

Correlations of the variables with the peer ranking within each DAFSC group and in all combined are shown in Table 5. Variables which did not correlate in any group were omitted. Although the correlations are low, some attain statistical significance within each group. It is of interest that variables which correlate in one DAFSC group tend not to correlate in another. Most relationships involving UPT grades are positive for fighter pilots and negative for others, possibly reflecting a fighter orientation in UPT.

Higher correlations may be expected from weighted combinations of the variables and their interactions. The AID model offers a convenient way to compute such multiple correlations. Use of this model requires that all variables except the peer ranking be converted to categorical form. A split is then made between categories of some variable, dividing the sample into two groups. The split location is chosen by a computer to maximize the amount of variance in the peer ranking which can be accounted for by the predictors at that stage. This in effect takes all possible interactions of predictors into account. The process is repeated with another split determined in the same manner until a stop criterion based on group size and increase in explained variance is reached.

*Table 5.* **Relationships between Significant Predictors and Performance**

| Variable | Transport ($N$=115) | Fighter ($N$=243) | Recce ($N$=34) | All ($N$=392) |
|---|---|---|---|---|
| **AFOQT Variables** | | | | |
| AFOQT Verbal Percentile | −.09 | −.14* | .28 | −.08 |
| AFOQT Quantitative Percentile | −.02 | −.11 | .39* | −.07 |
| **UPT Variables** | | | | |
| Rank Cadet on Entry into UPT | −.01 | .13* | −.05 | .06 |
| Principles of Flight, Primary | −.33* | −.01 | −.13 | −.10* |
| Aviation Physiology, Primary | −.22* | .03 | −.20 | −.06 |
| Engineering and Systems, Primary | −.21* | .01 | .04 | −.11* |
| Flight Instruments, Primary | −.28* | −.03 | −.26 | −.14* |
| Navigation, Primary | −.19* | −.09 | −.13 | −.10* |
| Contact Flying, Primary | −.05 | .15* | −.22 | .05 |
| Instrument Flying, Primary | −.18* | .00 | .00 | −.08 |
| T-33 Basic Aircraft | .15 | .11 | .25 | .24* |
| T-38 Basic Aircraft | −.18* | −.11 | −.25 | −.25* |
| Instrument Procedures, Basic | −.08 | .05 | −.06 | −.11* |
| Academic Average, Basic | −.20* | .06 | −.03 | −.02 |
| Navigation, Basic | −.20* | .06 | −.03 | .00 |
| Contact Flying, Basic | −.18* | .14* | −.06 | .00 |
| Instrument Flying, Basic | −.21* | .06 | −.07 | −.01 |
| Formation Flying, Basic | −.08 | .18* | −.09 | −.01 |
| Flying Training Grade | −.22* | .20* | .02 | .03 |
| **Biographical Variables** | | | | |
| Source of Commission Officer Training School | −.19* | −.11 | .00 | −.17* |
| Source of Commission Other | .08 | .07 | .08 | .10* |
| Years of Education beyond High School at End of UPT | −.10 | −.16* | −.24 | −.11* |
| Favorable Attitude Toward First Assignment after UPT | .00 | .11 | .24 | .10* |
| Academic Field Business | −.16 | −.01 | −.36* | −.05 |
| Academic Field Education | −.03 | −.13* | .19 | −.12* |
| Academic Field Unspecified | .03 | .14* | .22 | .12* |

*Significant at .05 level.

Table 6 shows the multiple correlations obtained from the AID analyses. Many of the correlations involve a large number of variables in relation to the sample sizes and are therefore inflated. Analyses leading to extreme inflation were omitted. There is some inflation in the cross-validation coefficients because the cross-validations were of the splits rather than the weights. In these circumstances, it appeared best to impose a stringent significance level of .001 on the correlations. Nearly all correlations are nevertheless significant. The inflation problem is minimal where only AFOQT variables are used as predictors, and here the correlations range as high as .54 on cross-validation. The peer ranking appears to be predictable within DAFSC groups and even within more homogeneous groups defined by specific type of aircraft flown. No cases were withheld for cross-validation in the groups defined by type of aircraft, however, and analyses involving only AFOQT variables were not performed.

*Table 6.* **Multiple Prediction of Performance by AID Analysis**

| Pilot Group | Category of Variable | | |
|---|---|---|---|
| | AFOQT | All Pre-UPT | All Pre-UPT Plus UPT |
| Transport, Subsample A[1] | .87* | .88* | |
| Transport, Subsample B[1] | .54* | .50* | |
| Fighter, Subsample A[2] | .72* | .80* | .88* |
| Fighter, Subsample B[2] | .41* | .45* | .47* |
| Recce, Subsample A[3] | .92* | | |
| Recce, Subsample B[3] | .36 | | |
| C-130, Subsamples A and B[4] | | .66* | .79* |
| F-4, Subsamples A and B[5] | | .65* | .76* |
| All, Subsample A[6] | .80* | .83* | .88* |

[1] $N$ = 115 in subsample A and 115 in subsample B (cross-validation subsample).

[2] $N$ = 243 in subsample A and 242 in subsample B (cross-validation subsample).

[3] $N$ = 34 in subsample A and 35 in subsample B (cross-validation subsample).

[4] $N$ = 196 in subsamples A and B combined.

[5] $N$ = 286 in subsamples A and B combined.

[6] $N$ = 392 in subsample A.

*Significant at .001 level.

## XII. MULTIPLE DISCRIMINANT ANALYSIS

The part of this study which directly addresses the TAC problem is the multiple discriminant analysis. A discriminant analysis is a procedure for estimating the position of an individual on a line that best separates classes or groups (Cooley & Lohnes, 1962). The simplest form of discriminant analysis can be represented geometrically as in Figure 1.

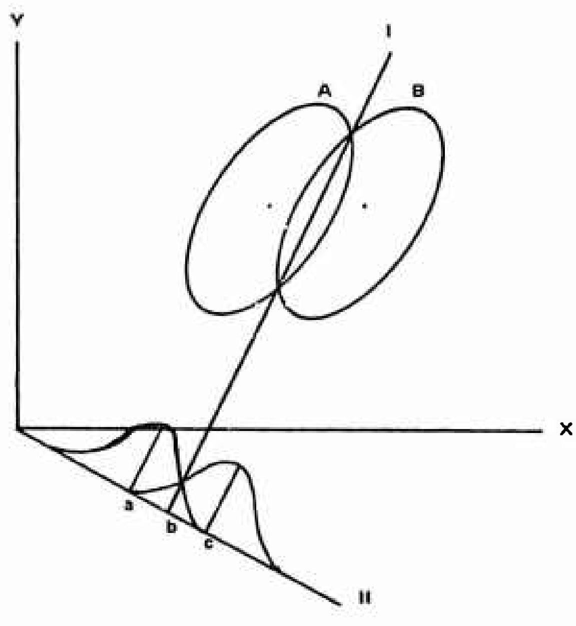


*Fig. 1.* **Geometric representation of discriminant analysis. (Adapted from Cooley & Lohnes, 1962.)**

Figure 1 depicts two groups, A and B, which differ in their means on two measures, X and Y. The groups might be transport and fighter pilots, and the measures the Basic Academic Average and Flying Training Grade. For each individual, a point can be plotted which represents his scores on both measures. The point is plotted on a surface defined by axes X and Y. Such points will tend to fall in two clusters, at the centers of which are the group means. One can then draw two ellipses which enclose all points for the two groups separately, or all points except for extreme deviants. The elliptical shape indicates a degree of correlation between the measures. The points where the ellipses intersect define a line I. If a second line, II, is drawn perpendicular to I, and if the points on the surface are projected onto II, the overlap between the groups will be minimized. Point b, at the intersection of the lines, divides line II into segments representing probable membership in group A and probable membership in group B. Points a and c are the projected means of the groups and are known as centroids.

An individual's position on line II defines his discriminant score and is a linear function of his values on measures X and Y. When an individual of unknown group membership appears, such as a new UPT graduate, a discriminant score is computed for him from X and Y, and he is classified according to the segment of the line in which his score falls.

The number of groups can be increased, as by the addition of a recce group, and the number of measures can be increased, as by the addition of other variables from Table 2. There may then emerge more than one discriminant function. Geometric representation breaks down in these more complex situations, but the basic principle of discriminant analysis remains unchanged.

## XIII. DISCRIMINATION AMONG ABOVE AVERAGE PILOTS

Several different but highly overlapping groups of variables were used to find functions for discriminating among above average pilots in the three DAFSC groups. The choice of variables was guided initially by the expectation that good predictors for one group might discriminate between that group and another. At a later stage, the correlations with discriminant functions became available as an additional guide. Effective discrimination was obtained from nearly all sets of variables tried, but most sets were not efficient in that they contained variables which contributed nothing to the discrimination. It was not possible to discriminate from the AFOQT variables alone.

An efficient discriminant function was developed from ten AFOQT and UPT variables. It was the first of two functions to emerge from the particular analysis. The second function was not statistically significant and can be ignored. It had an associated probability value of .10. This is an advantageous outcome from the point of view of ease of application. In determining how to assign a new pilot, it is necessary to compute only one discriminant score.

The properties of the significant discriminant function are summarized in Table 7. The table shows for each variable the weight applied to it to yield the function, the mean of each variable in the three above average pilots groups, and the correlation of each variable with the function. The final column shows the probability values associated with each variable, the first function, and Wilks' lambda. Wilks' lambda, with a numercial value of .802, measures the over-all discriminating power of the system. The associated probability value of .00 indicates that the chances of producing group differences at random as large as those actually obtained are essentially zero. The centroids are also shown in the table.

## XIV. DETERMINATION OF A PILOT ASSIGNMENT

The discriminant analysis summarized in Table 7 can be used to assign a new pilot to a DAFSC group. The procedure is as follows:

1. Insert his scores on the ten variables in the following equation:

$$Y_d = -(.0645)X_1 + (.2388)X_2 + (.1638)X_3$$
$$-(.0236)X_4 + (.1829)X_5 + (.088\text{[illegible]})X_6 + (.3874)X_7$$
$$-(.0110)X_8 + (.3890)X_9 + (.7540)X_{10}$$

where $Y_d$ is the discriminant score, and the subscripted Xs are the values of the variables in Table 7 which bear the numbers of the subscripts.

2. Multiply each score by its proper weight and sum the products algebraically to obtain the discriminant score. Retain four decimal places.

3. If the discriminant score is less than 154.69199, the assignment should be to the transport DAFSC group. If the discriminant score

*Table 7.* **Multiple Discriminant Analysis of Above Average Pilots**

*(N = 119 transport pilots, 246 fighter pilots, and 39 recce pilots.)*

| Variable | Weight | Trans-port Mean | Fighter Mean | Recce Mean | Corre-lation with Function | Proba-bility |
|---|---|---|---|---|---|---|
| 1. AFOQT Verbal Percentile | −.06 | 59.05 | 52.82 | 63.33 | −.32* | .02 |
| 2. Principles of Flight, Primary | .24 | 86.50 | 88.11 | 87.05 | .50* | .00 |
| 3. Flight Instruments, Primary | −.16 | 86.41 | 87.80 | 87.87 | .36* | .00 |
| 4. Navigation, Primary | −.02 | 87.07 | 87.96 | 86.79 | .36* | .01 |
| 5. Contact Flying, Primary | .18 | 86.34 | 87.96 | 87.72 | .52* | .00 |
| 6. Instrument Flying, Primary | .09 | 86.46 | 88.28 | 87.46 | .52* | .00 |
| 7. Academic Average, Basic | .39 | 85.76 | 87.39 | 86.23 | .51* | .00 |
| 8. Instrument Flying, Basic | −.01 | 86.24 | 88.16 | 86.56 | .62* | .00 |
| 9. Formation Flying, Basic | .39 | 86.13 | 88.38 | 86.41 | .71* | .00 |
| 10. Flying Training Grade | .75 | 85.26 | 87.84 | 86.23 | .85* | .00 |
| Function 1 | | | | | | .00 |
| Wilks' Lambda | | | | | | .00 |
| Centroids | | 154.19 | 158.62 | 155.17 | | |

*Significant at .05 level.

is more than 154.69199 but less than 156.84134, the assignment should be to the recce DAFSC group. If the discriminant score is more than 156.84134, the assignment should be to the fighter DAFSC group.

If only a small number of pilots are to be assigned at one time, the computations can be performed readily with a desk calculator and locally prepared worksheets. The results should be audited, preferably by a second person. If there are many pilots to be assigned, it may be desirable to computerize the routine. Very little computer time would be required for any group that could realistically be expected.

The values 154.69199 and 156.84134 are equivalents of point b in Figure 1. The overlapping distributions which they maximally separate are shown in Table 8. Because the distributions have nearly the same dispersions, the values are close to the midpoints between adjacent centroids. The values could be established elsewhere between centroids if necessary to meet any overriding assignment quotas. While the system provides for assignments to only three DAFSC groups, it might reasonably be used in other assignment situations to determine which pilots should be assigned to fighter type aircraft and which to heavy multi-engine aircraft.

*Table 8.* **Distributions of Discriminant Scores for Above Average Pilots**

| | Transport | | Fighter | | Recce | |
|---|---|---|---|---|---|---|
| Discriminant Score | *N* | Percent | *N* | Percent | *N* | Percent |
| 138.00 - 141.99 | 1 | 0.8 | 0 | 0.0 | 1 | 2.6 |
| 142.00 - 145.99 | 5 | 4.2 | 1 | 0.4 | 1 | 2.6 |
| 146.00 - 149.99 | 11 | 9.2 | 6 | 2.4 | 1 | 2.6 |
| 150.00 - 153.99 | 34 | 28.6 | 22 | 8.9 | 10 | 25.6 |
| 154.00 - 157.99 | 50 | 42.0 | 81 | 32.9 | 17 | 43.6 |
| 158.00 - 161-99 | 15 | 12.6 | 88 | 35.8 | 9 | 23.1 |
| 162.00 - 165.99 | 2 | 1.7 | 31 | 12.6 | 0 | 0.0 |
| 166.00 - 169.99 | 1 | 0.8 | 12 | 4.9 | 0 | 0.0 |
| 170.00 - 173.99 | 0 | 0.0 | 5 | 2.0 | 0 | 0.0 |
| Total | 119 | 99.9 | 246 | 99.9 | 39 | 100.1 |
| Mean | 154.19 | | 158.62 | | 155.17 | |
| Standard Deviation | 4.55 | | 4.61 | | 4.46 | |

## XV. COMPARISON OF ACTUAL AND OPTIMAL ASSIGNMENTS

If there are advantages in this assignment system, they should be reflected in differences between pilots who are actually assigned compatibly with it and those not so assigned. One manifestation should be a larger proportion of compatibly assigned pilots with above average peer rankings than with average or below average rankings. Table 9 shows that this expectation is supported in every DAFSC group. Among above average pilots, 58 percent are compatibly assigned. Among other pilots, only 43 percent are compatibly assigned. There should also be more above average pilots with compatible than incompatible assignments and fewer other pilots with compatible than incompatible assignments. The table confirms these expectations for actual transport and fighter pilots and partially confirms them for recees.

*Table 9.* **Actual and Optimal Assignments for Six Pilot Groups**

| Actual | Optimal | Above Average Pilots | | Average and Below Average Pilots | |
|---|---|---|---|---|---|
| | | *N* | Percent | *N* | Percent |
| Transport | Transport | 66 | 55.46 | 43 | 38.74 |
| Transport | Fighter | 28 | 23.53 | 47 | 42.34 |
| Transport | Recce | 25 | 21.01 | 21 | 18.92 |
| Fighter | Fighter | 157 | 63.82 | 115 | 48.12 |
| Fighter | Recce | 52 | 21.14 | 62 | 25.94 |
| Fighter | Transport | 37 | 15.04 | 62 | 25.94 |
| Recce | Recce | 11 | 28.21 | 7 | 23.33 |
| Recce | Transport | 16 | 41.03 | 10 | 33.33 |
| Recce | Fighter | 12 | 30.77 | 13 | 43.33 |

If all cases in the total sample were apportioned among the three DAFSC groups as in the TAC population, there would be 264 transport pilots, 458 fighter pilots, and 62 recce pilots. These proportions presumably express the requirements of the Air Force. If all cases in the sample were assigned according to the optimal system, there would be 234 transport pilots, 372 fighter pilots, and 178 recce pilots. Hence, all pilots optimally assignable to transport and fighter DAFSC groups could actually be assigned in this way. Also, all recce assignments could be filled by pilots optimally assignable as recces. This leaves 116 optimal recces who must be assigned elsewhere. These cases amount to 15 percent of the total sample. In the sample as actually constituted, 49 percent are not optimally assigned.

The assignment system, moreover, provides guidance for assignment of the 116 cases. One can assign to transports those whose discriminant scores deviate most toward transports and to fighters those who deviate most toward fighters until transport and fighter quotas are filled. The remaining 62 will be recces. In this way, all cases can be assigned with regard for their discriminant scores. If it is required that those assigned as recces be closest to the recce centroid, one can still assign the others with regard for their discriminant scores except for 27 cases. These 27 make up only 3.4 percent of the total sample

## XVI. VALIDATION OF THE ASSIGNMENT SYSTEM

A thorough validation of this system, including cross-validation, requires a large scale study. This could be conducted within TAC while all actual assignments are made as at present. For each pilot assigned to TAC, a discriminant score could be computed and recorded as an experimental datum. Ratings or other performance measures could ultimately be related to discrepancies between actual and optimal assignments. The system should be proposed for operational use only when the data have accumulated convincingly in its favor.

A preliminary validation can be performed with existing data. It is assumed that, throughout a wide range of pilot competence, pilots will perform best in assignments compatible with their aptitudes. If the system is valid, the average competence of pilots who are assigned compatibly should be greater than that of pilots not so assigned, and the difference should appear in the form of higher mean peer rankings.

To test this hypothesis, all compatibly assigned pilots within each DAFSC group were combined, regardless of their peer ranking, and the means and standard deviations of their peer ranking distributions were computed. Incompatibly assigned pilots were treated similarly. The results are shown in Table 10.

*Table 10.* **Performance of Compatibly and Incompatibly Assigned Pilots**

| Actual | N | Mean | SD | t |
|---|---|---|---|---|
| Compatible Transport | 109 | 52.32 | 7.15 | 2.95* |
| Incompatible Transport | 121 | 49.49 | 7.36 | |
| Compatible Fighter | 272 | 51.77 | 7.45 | 3.40* |
| Incompatible Fighter | 213 | 49.49 | 7.27 | |
| Compatible Recce | 18 | 54.00 | 8.00 | 1.78 |
| Incompatible Recce | 51 | 50.24 | 7.44 | |
| Compatible Pilot | 399 | 52.03 | 7.41 | 4.63* |
| Incompatible Pilot | 385 | 49.59 | 7.32 | |

*Statistically significant at .05 level or beyond.

The table shows that within each DAFSC group there is a higher mean peer ranking for compatibly assigned pilots than for others. This is also true when the groups are collapsed and all compatibly and incompatibly assigned pilots are compared at once, as in the final section of the table. Moreover, the recce mean difference, which is the largest of all, is the only one failing statistical significance as determined by the *t* test.

The practical importance of these differences must be judged on bases other than statistical tests. Since flying safety, job satisfaction, mission effectiveness, and training efficiency are all presumed to be influenced by appropriateness of assignment, any statistically significant differences should probably be taken seriously.

## XVII. CONCLUSIONS

This research supports the following conclusions:

1. Pilots generally appear able to agree beyond chance expectations in their independent ranking of other pilots in terms of over-all competence.

2. There are several sets of variables whose values are known by the end of UPT which predict performance of transport, fighter, and recce pilots, and of all combined. At least two sets predict performance in C-130 and F-4 aircraft.

3. Variables which best predict performance in one pilot specialty tend not to be the ones which best predict performance in a different specialty.

4. UPT grades tend to be positively related to the performance of fighter pilots, but negatively related to the performance of transport and recce pilots.

5. Several sets of variables whose values are known by the end of UPT can be used to predict for a pilot that he is optimally assignable to the transport, fighter, or recce DAFSC group. An efficient system exists for making such assignments from ten AFOQT and UPT variables which define a single significant discriminant function.

6. A majority of pilots with above average peer rankings are assigned at present in accordance with their optimal assignments. The proportion of such compatible assignments is smaller for pilots with average and below average peer rankings.

7. Within the constraint of Air Force quotas, it is possible to assign optimally 85 percent of the pilots in this study and to assign the remainder with some consideration for their optimal assignments. Only 51 percent are now optimally assigned.

8. Pilots whose actual assignments are compatible with their optimal assignments are better pilots, as measured by peer rankings, than pilots not compatibly assigned. The differences are statistically significant except in the case of recces.

## REFERENCES


Cooley, W.W., & Lohnes, P.R. *Multivariate procedures for the behavioral sciences.* New York: Wiley, 1962.

DuBois, P.H. (Ed.) *The classification program.* Army Air Forces Avaiation Psychology Research Report No. 2. Washington, D.C.: United States Government Printing Office, 1947.

Flanagan, J.C. (Ed.) *The aviation psychology program in the Army Air Forces.* Army Air Forces Aviation Psychology Research Report No. 1. Washington, D.C.: United States Government Printing Office, 1948.

Hulin, C.L., & Alvares, K.M. *An evaluation of three possible explanations of the temporal decay in predicting pilot proficiency.* AFHRL-TR-71-5, AD-731 191. Williams AFB, Ariz.: Flying Training Division, Air Force Human Resources Laboratory, February 1971.

Kendall, M.G. *The advanced theory of statistics.* Vol. 1. London: Charles Griffin and Co., Ltd., 1945.

Koplyay, J.B., Gott, C.D., & Elton, J.H. *Automatic Interaction Detector-Version 4 (AID-4) Reference Manual.* AFHRL-TR-73-17, AD-773 803. Lackland AFB, Tex.: Personnel Research Division, Air Force Human Resources Laboratory, October 1973.

Miller, R.E. *Interpretation and utilization of scores on the Air Force Officer Qualifying Test.* AFHRL-TR-69-103, AD 691 001. Lackland AFB, Tex.: Personnel Research Division, Air Force Human Resources Laboratory, May 1969.

Peterson, F.E., Booth, R.F., Lane, N.E., & Ambler, R.K. *Predicting success in naval flight officer training.* NAMI-996. Pensacola, Fla.: Naval Aerospace Medical Institute, Naval Aerospace Medical Center, February 1967.

Sonquist, J.A., & Morgan, J.N. *The detection of interaction effects.* University of Michigan Survey Research Center Monograph No. 35. Ann Arbor, Mich.: University of Michigan Press, 1964.

Torrance, E.P. *The development of a preliminary life experience inventory for the study of fighter interceptor pilot combat effectiveness.* AFPTRC-TR-54-89, AD-64 759. Lackland AFB, Tex.: Air Force Personnel and Training Research Center, December 1954.